

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 15, 2002, 02:56:03 ; Search time 3106.77 Seconds
(without alignments)
23337.777 Million cell updates/sec

Title: US-09-652-292-1

Perfect score: 4395

Sequence: 1 gagggggtccttgcaggcc.....attattgtataaaaaaa 4395

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.om.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.sy.*

28: em.un.*

29: em.vi.*

30: em.htgo_hum.*

31: em.htgo_inv.*

32: em.htgo_rtd.*

33: em.htg_hum.*

34: em.htg_inv.*

35: em.htg_rtd.*

36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	4395	100.0	4396	9	AF248053	AF248053 Homo sapi
2	4105.4	93.4	4167	9	AF321240	AF321240 Homo sapi
3	4064.4	92.5	4126	9	HS28H201	AL137188 Novel hum
4	2531.4	57.6	127418	9	HS28H20	AL031055 Human DNA
5	795	18.1	200125	2	AL591064	AL591064 Mus muscu
6	795	18.1	243075	2	AL591805	AL591805 Mus muscu
7	384.6	8.8	411	11	G31518	G31518 human STS S
8	317.4	7.2	385	6	AX150151	AX150151 Sequence
9	208.6	4.7	4521	9	AB056798	AB056798 Macaca fa
10	207.2	4.7	173874	2	AC068011	AC068011 Homo sapi
11	201	4.6	174278	2	AC016330	AC016330 Homo sapi
12	199.8	4.5	173681	2	AC009502	AC009502 Homo sapi
13	199.8	4.5	187336	9	AL158822	AL158822 Human DNA
14	199.8	4.5	85378	2	AL356577	AL356577 Homo sapi
15	197.2	4.5	98360	9	HS247C2	AL049713 Human DNA
16	197.2	4.5	117853	9	AC007030	AC007030 Homo sapi
17	196.2	4.5	107467	9	HS95883	293023 Homo sapien
18	196	4.5	160421	2	AC025005	AC025005 Homo sapi
19	195.8	4.5	225022	2	AL358779	AL358779 Homo sapi
20	195.8	4.5	92280	2	AC092769	AC092769 Homo sapi
21	195.6	4.4	167178	2	AC025672	AC025672 Homo sapi
22	195.4	4.4	156776	9	AC011443	AC011443 Homo sapi
23	194.8	4.4	130615	9	AC010395	AC010395 Homo sapi
24	194.6	4.4	167275	9	AC008782	AC008782 Homo sapi
25	194.6	4.4	149411	2	AC019029	AC019029 Homo sapi
26	193	4.4	160625	9	AC009311	AC009311 Homo sapi
27	193	4.4	152408	9	HS191118	AL024507 Human DNA
28	192.8	4.4	176641	2	AL450043	AL450043 Homo sapi
29	192.6	4.4	184877	2	AL451000	AL451000 Homo sapi
30	192.6	4.4	194750	2	AC092961	AC092961 Homo sapi
31	192.4	4.4	196267	2	AC026603	AC026603 Homo sapi
32	192.4	4.4	294754	2	AC055741	AC055741 Homo sapi
33	192.4	4.4	89881	9	AL160264	AL160264 Human DNA
34	192	4.4	245077	9	AF002997	AF002997 Homo sapi
35	192	4.4	22289	2	AL450024	AL450024 Homo sapi
36	191.8	4.4	56330	2	AL353694	Continuation (4 of
37	191.6	4.4	89122	9	AC007199	AC007199 Homo sapi
38	191.6	4.4	170923	2	AL136380	AL136380 Homo sapi
39	191.6	4.4	171462	2	AL359473	AL359473 Homo sapi
40	191.6	4.4	120689	9	CNS01RGS	AL159191 Human chr
41	191.4	4.4	147430	2	AC092066	AC092066 Homo sapi
42	191.4	4.4	148973	2	CNS01DUZ	AL133374 Homo sapi
43	191.4	4.4	167294	2	AC013814	AC013814 Homo sapi
44	191.4	4.4	126801	9	AC004849	AC004849 Homo sapi
45	191.2	4.4				

ALIGNMENTS

RESULT 1

AF248053

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AF248053 4396 bp mRNA PRI 12-APR-2001
Homo sapiens glucose transporter (GLUT10) mRNA, complete cds.

AF248053.1 GI:13603726

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4396)

Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychalecky,J.C.,

Dawson,P.A. and Bowden,D.W.

GLUT10: A novel glucose transporter in the type 2 diabetes linked

region of chromosome 20q12-13.1

Unpublished

2 (bases 1 to 4396)

Fossey,S.C., Mihic,S.J., Craddock,A.L., Mychalecky,J.C.,

Dawson,P.A. and Bowden,D.W.

TITLE Direct Submission
JOURNAL Submitted (22-MAR-2000) Molecular Genetics, Wake Forest University,
Medical Center Boulevard, Winston-Salem, NC 27106, USA
FEATURES Location/Qualifiers
Source i. .4396
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q12-q13.1"
i. .4396
/gene="GLUT10"
251..1876
/gene="GLUT10"
/codon_start=1
/product="glucose transporter"
/protein_id="AAK31911.1"
/db_xref="GI:13603727"
/translation="MGHSPVPLCASVLLGLTFYELAVISGALLPLQLDFGLSC
LEQEFVGSLLLCALLSLVGGFLIDCYGRKQALLGNLVLGSLALYALNLAGT
LGRAVVGPAISLSMACCIYVSELVGRQVLSVLEAGITVGLISYALNLAGT
PKWRHMFQWATAPAVLSLLFLPAGTDETHKDLIPLQGEAPKLPGRPRYSF
LDLFRANDNRGRTTVGLGLVLFQQLTGQPNVLCYASTIFSSVGFHGGSSAVLASVL
GAVKAATIMGLVDRAGRALLAGCALMALSVSIGLVSFAPVMDSPSCLAVPN
ATGQGLPDGSLDSSLPPIPTNEDOREPILSTAKTKPHRSPGDPSPAPRALUS
SALPGPLPARGHALLRWTLICLVFVSFSGFGVPTWLVLSIYPIVEIRGFAF
CNSFNAAANLFSLSLDLIGTGLSWTFELGLTAVLGLGFIYLFVPETRKQSLAEI
DOOFKRRPTLSFGHONSTGIPYSRIEISAA"

BASE COUNT 939 a 1139 c 1102 g 1216 t
ORIGIN

Query Match 100.0%; Score 4395; DB 9; Length 4396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4395; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gagggggtcttcgagggcctggcgccggggcggtgctgagctgagctccctccctccg 60
DB 2 GAGGGGTCTTGGCCAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 61
QY 61 cctcagggcctggggcctggggcctggggcctggggcctggggcctggggcctgggg 120
DB 62 CCTCAGGCTCGGGGGCTGGCTGGCCGACGTGGCGCTGGCGGGCTGGCGGGGAGGG 121
QY 121 caggggcaggaggacagaggcggggcgccggcggcggcggcggcggcggcggcgg 180
DB 122 CAGGCGAGGAGGACAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 181
QY 181 ggggactcgggggggagtgcgccggcggcggcggcggcggcggcggcggcggcgg 240
DB 182 GGGGACTCGGGGGGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTC 241
QY 241 ccgctccatggccactcccaactgctgctgctgctgctgctgctgctgctgctgctg 300
DB 242 CCGCTCGCATGGGGCACTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
QY 301 ggtggcctgaccttgggttaagactggcagtcataatcagtgccctgctgctgctgctg 360
DB 302 GGTGGCTGACCTTTGTTATGAAGTGGCAGTCATATCAGGTGCGCTGCTGCTGCTGCTG 361
QY 361 ctgactttgggctaagctgctggagcaggagttccctgggtgggagcctgctgctgctg 420
DB 362 CTTGACTTTGGGCTAAGCTGCTGGAGCAGAGTTCCTGGTGGGGCAGCCTGCTCCTGGGG 421
QY 421 gctcctcgcctccctcgggttggtggttctctcctcctcctcctcctcctcctcctc 480
DB 422 GCTCTCTCGCCCTCGGCTGGTGGTGGCTTCTCATTTGACTGCTATGAGGAGGAAGGCC 481
QY 481 atcctcgggagcaacttgggtgctgctgagcagcagcctgacccctgggctgctgctg 540
DB 482 ATCCTCGGAGCAACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 541
QY 541 ctggcctgggtgctcctggggcggcgtggtggtggtggtggtggtggtggtggtggtg 600
|||||

DB 542 CTGGCCTGGCTGGTCTGCTGGCGCGCGCTGTGTTGGCTTGCCATTTCCTCTCCCTCCAG 601
QY 601 gcttctgtatatacgtgtcagagctggtggggccacgagcggggagtgctgctgctgct 660
DB 602 GCTTGCTGTATCTACGTGTGTCAGAGCTGGTGGGGCCACGGCAGCGGGAGTCTGGTGTC 661
QY 661 ctctatgaggcaggcatcacctgggcatcctgctcctctctctctctcctcctcctcctg 720
DB 662 CTCTATGAGGAGGACATCACCGTGGGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 721
QY 721 gctggtacccctcggggatggaggacatgttcggtcgggtgggcccacacacacacac 840
DB 722 GCTGGTACCCCTCGGGATGGAGGCATGTTCCGCTGGCCACTGCACCTGCTGCTGCTG 781
QY 781 caatccctcagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 840
DB 782 CAATCCCTTCAGCCT 841
QY 841 ctcatccactccaggaggtagagcccccagcctggggccgggggggggggggggggggg 900
DB 842 CTCTATCCACTCCAGGAGGTGAGGCCCCCAAGCTGGGGCCGGGGGAGGCCACGGTAC 901
QY 901 ttcttgacctcttcagggcacgcgataacatgagggcggcggcggcggcggcggcggg 960
DB 902 TTTCTGGACCTCTTCAGGGCAGCGATACATGCCAGGCCGGACACAGTGGGCTGGGG 961
QY 961 ctgggtctctccagcaactaacaggggcagcccaacgtgctgctgctgctgctcctcct 1020
DB 962 CTGGTCTCTTCCAGCAACTAACAGGGCAGCCCAAGCTGCTGTGCTATGCTCCACCATC 1021
QY 1021 ttcagctccgcttgggtttccatggggatcctcagcggctgctgctgctgctgctgctg 1080
DB 1022 TTCAGTCCGCTGGTGTTCATGGGGGATCCCTCAGCGCTGCTGGCCTCTGCTGGGCT 1081
QY 1081 gcaagtgaagtgggcagctacccctgacgcccctggggcgtggtgagcgtgacggcgag 1140
DB 1082 GCAGTGAAGTGGCAGTACCTCAGCCGCATGCCCATGGGCTGGTGGACCGTGCAGGCC 1141
QY 1141 gctctgtttagctggtgctgctgctgctgctgctgctgctgctgctgctgctgctg 1200
DB 1142 GCTCTGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1201
QY 1201 agctttgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1260
DB 1202 AGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
QY 1261 cagcaggcctcctcctggagactcctgctgctgctgctgctgctgctgctgctgctg 1320
DB 1262 CAGCAGGGCTCTCCTGGAGACTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1321
QY 1321 agggacaaatgagacacaaaggagcacaatttctcactgctaagaaacacaaagcccat 1380
DB 1322 AGGACCAATGAGACCAAAAGGGAGCAATCTTGTCCACTGCTAAAGAAACCAAGCC 1381
QY 1381 cccagatctggagacccctcagccctcctcctgctgctgctgctgctgctgctgctgctg 1440
DB 1382 CCCAGATCTGGAGACCCCTCAGCCCTCTCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1441
QY 1441 cccctcctgcccgcctggggcagtcgctgctgctgctgctgctgctgctgctgctg 1500
DB 1442 CCCCCTCTGCCGCTCGGGGGATGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
QY 1501 gcttttgcagtgcttctccttgggttggggcagtgacctgggttctcctcagcag 1560
DB 1502 GTCCTTTGTCAGTGCCT 1561
QY 1561 atctacacctgtggagatacagaggaagagccttcgctcctcctcctcctcctcctc 1620
DB 1562 ATCTACCTCTGGAGATACGAGGAAGAGCCTTCGCCCTTCCTGCAACAGCTTCACT 1621
QY 1621 gcaaacctcttctcagcctcctcctcctcctcctcctcctcctcctcctcctcctc 1680
DB 1622 GCCAACCTCTTCTATCAGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1681

Qy	1681	accltccctgtctctacggaactgaccgctgtcctctggcctgggcttcactctattttgtt	1740
Db	1682	ACCTTCCTGCTCTACGGACTGACCGCTGCTCTCGGCTGGCTTCATCTATTATTGTT	1741
Qy	1741	cctgaacaaagaagccagtcgttggcagagatagaccagacagttccagaagagacggttc	1800
Db	1742	CCTGAACAAAGGCCAGTCTGTTGCGAGAGATAGACAGCAGTTTCCAGAAGACGGTTC	1801
Qy	1801	accctagctttggccacagcagaactccactggcatcccgctacagccgcatcgagatc	1860
Db	1802	ACCCTGAGCTTTGGCCACAGGCAGAACTCCACTGGCATCCCGTACAGCCGATCAGATC	1861
Qy	1861	tctgggctccctcctgaggaatccgtctgctggaatctctggaactgtggtcttggcagacc	1920
Db	1862	TCTGGCGCTCTCTGAGGAATCGCTCTGCTCGGAATTCGTGGAACCTGGCTTTGGCAGACC	1921
Qy	1921	atctccagcatcctgtctctctaggccccagagacaaagttccagctggtctctttggagct	1980
Db	1922	ATCTCCAGCATCCTGCTTCTTAGGCCCCAGAGCACAAAGTTCCAGCTGCTCTTTGGGAGT	1981
Qy	1981	ggcccttgcctcccaaaagtgtgtcttctgctggggtaaaaaggaagaaagtttgagaa	2040
Db	1982	GGCCCTTGCCCCAAAGGTGTTGCTTTGCTGGGTAAAAAGGATGAAAGTTTGAGAA	2041
Qy	2041	tgcccaattcttcaatcttttgggttccaggccctgaagttctcttgaggatctctagttcatgc	2100
Db	2042	TGCCCAATTCTTCATTTTGGGTTTCAGGCCCTCAGAGTTCTTGAGGATCTAGTTTTCATGC	2101
Qy	2101	ctgggttccccaatgtaacttggacaattttgcagatttttataagaagaatatctcatgaa	2160
Db	2102	CTGGTTCCTCCCATTGACTTGGACATTTTTCAGTTTATAGAAGAAATATTCTATGAA	2161
Qy	2161	gtcttgttgcctcccatggatttttttcaagaatactcaggggtaccaatccgggcaggag	2220
Db	2162	GTCTTTGTGCCCACTGGATTTTTTTTCAAAGAAATCTCAGGGGTACCAATCGGGCAGGAG	2221
Qy	2221	gttttccccgatatcccccataatccaaatgaggaatcatcttttctaactctctttt	2280
Db	2222	GTTTTTCCTCCGATATCACCCCTAAATTCCAATGAGGATATCATCTTTCTAATCTCTTTT	2281
Qy	2281	tcaacttgccttgggacatttccggaaggggaagctctcttttctactcttatctttt	2340
Db	2282	TCAACTGGCTGGGACATTTTCGGAAGGGGAAGCTCTCTTTTTTTTATCTATCTTTTTT	2341
Qy	2341	tttttggagtggaagtcatactgtgtgccccagctggcctgatacttggcctcactlgcaac	2400
Db	2342	TTTTTTTGAGTGGAGTCTCATCTGTTGCCAGGCTGGCCTGATCTTGGCTCACCTGCAAC	2401
Qy	2401	ctccacttctgggttccnaagcgattctctgctcagctcctctaagtagctgggattaca	2460
Db	2402	CTCCACTCTCTGGGTTCAAGCGATTCTCTGCCTCAGCCCTTAAGTAGCTGGGATTACA	2461
Qy	2461	ggcgtgtgccaaccacccagctaatatttttttagcagagatgggggttctcactgtgttg	2520
Db	2462	GGCGCTGCCACCACACCCAGCTAATTTATTTTATAGCAGAGATCGGGTTTCTACTCTGTTG	2521
Qy	2521	gccaggtgtgtgtgaactcctgagctcaagtgatccacccactcagcctccccagagtg	2580
Db	2522	GCCAGGCTGGTGTGTAACCTCTGAGCTCAAGTGCATCCACCACCTCAGCCCTCCAGAGTG	2581
Qy	2581	ctaggattacagggccttttgactcttttatcttgagtttttatgacccctctaattctctt	2640
Db	2582	CTAGGATTACAGGCTTTTGACTCTTTTATCTGAGTTTTATTGACCCCTCTAAATCTCTT	2641
Qy	2641	accagaataatttatcttaccagcaactctgactcttttgaagggaggcctcactctcta	2700
Db	2642	ACCACGAATATTTATCTCTACACCAACTCTCACTCTTTTGACGGGAGCCCTCAGTTCTA	2701
Qy	2701	gtccttggctgtggtgtcaattgctgtaggaaatgaccacgggcctcagtttccccattt	2760
Db	2702	GTCTTTGGTCTGTGTGTGTATGCTGTAGSAAATGACCAGGCCCTCAGTTTCCCATTTT	2761

QY	2761	gtataatggaagcctgtaccagggtcaattctttaaagattctctcctgaactccagctgaactcg	2821
DB	2762	GTATAATGGGAAGCCTGTACCAGGTCATCTCTTAAGATTTCTCTCTGACTCCAGTGAAGCTGG	2821
QY	2821	aatctctaaatgctggtctcaggagctgtctccaggatagtgcaggatggcctttgcggaag	2880
DB	2822	AAATCTAAATGCTGGTCTAGGAGCTGTCTCCAGGATGGTGCAGGATGGCTTTGCGGAAG	2881
QY	2881	gagctgggtttggaggccaacaaactgcttgtcaatatagcctttgctctcttggcagcc	2940
DB	2882	GAGATGGGTTTGGAGGCCAACAAACCTGCTGTCAAATATTCGCTTTGGCTCTTTGGCAGCC	2941
QY	2941	ctgaaactctgaatgaatacaactccctgaacctcagttctctcatctgcagaaatggga	3000
DB	2942	CTTGAACCTTGAGTAAATAAACAACCTCCCTGAACCTCAGTTCCTCATCTGCAGAAATGGGA	3001
QY	3001	taattatgtccagggttatatttagaacctgtttcctttcaggagggtccccagctgggt	3060
DB	3002	TAAATATCTCCAGGGGTATATTTAGACCTGTTCCTTTCAGGAGGGTCCCCAGAGTGGT	3061
QY	3061	ccagggtctgggaattcttaacttatctcctaatccceaggctccctcctttggaccctgta	3120
DB	3062	CCAGGGCTGGGAATTTCTACTTATCCCTCATTTACCCAGGTCCTCTTTTGGACCTGTA	3121
QY	3121	aagggtcaagggtgaatcagatgaggagactgagcaagttagctatgaactcagatcatgtaa	3180
DB	3122	AAGGTCAGGCTGAATCAGATGGGGGACTGAGCAAGTAGCTATGACTGCAGATCATGTAA	3181
QY	3181	ggaaggactgacaagaagctcccagatgctggggagaatgaagagctaaaaatagatcct	3240
DB	3182	GGAAGGCACTGACAAGAAGCTCCCAGATGCTGGGGAGAAATCAAGAGCTAAAAATAGATCCT	3241
QY	3241	aggctcggagctgttgcatacctcagctgtgcacatactgggtcgtggcagagagcccccaagg	3300
DB	3242	AGGTCGTGGATGCTTGTGTCATCCATCGCTGCACATATGGGTGCTGGCAGAGCCCCCAAGG	3301
QY	3301	actctggccttcagattctcctatctctccattctcagatcttagatctgcttcccttgatccagtg	3360
DB	3302	ACTCTGGCCTTCGAGTTCTCCTATCTTTCTCCATTTCTAGATGCTTCCCTTTGTATCCAGTG	3361
QY	3361	atgctcggagctggctttgcgaagctgtgcagagctgggttgctacatatttccaggtttt	3420
DB	3362	ATGTCGTGGAGCTGGCTTTGCCAAGCTTGTGAGAGCTGGTGTCTACATTTTCAGGATTTT	3421
QY	3421	tacaagtgtgtaaacacacagccattataaaaaattaaatgatttaaatattataatgaata	3480
DB	3422	TACAAGTTGGTTAAACACAGCCATTATAAAAAATTAATGATTTAAATTTATAATTAAGTA	3481
QY	3481	aattacattaaaaacaaaaaattatactcaaaattactaaattacttaactaccctgta	3540
DB	3482	AAATTACATTAACACAAAAAATTTATACTCAAAATTCATTAAATTTTACTTACCTGTGA	3541
QY	3541	ctattatctgcttttgaggctattcttcacatagtaacctctatggaagaccttagggag	3600
DB	3542	CTAATATCTGTGCTTTTGAGGCTATTCTTACATAGTAGTAACCTCTTATGGAGACCTTAGGGGAG	3601
QY	3601	acaccggcatctctcctgattccccactcaatgacatcatgttagtctttggttgctt	3660
DB	3602	ACACCGCGCATCTCTCTGATTTCCCACTCAATGACATCATATGTTAGTCTTTTGGTTGCTT	3661
QY	3661	aactgactggggagtgatttttgtatcacaaagattagagaggactacacatcagggct	3720
DB	3662	AACCTGCTGGGGAGTGTTTTTGTATCAAAAGATTTAGAGAGGACTACACATCAGGGCT	3721
QY	3721	tgaattattgtttgtgatttcttagacctcagaacatgctgggtataaaatgtaaaatg	3780
DB	3722	TGATTTATGTTGTGATTTTCTAGACTTCAGAACATGCTGGATATAAATGTACGTAATG	3781
QY	3781	gaattaaaactttaaagatgctctgtttgttagccaatcactggtgatatagccacccaaa	3840
DB	3782	CAAAATTAACCTTTAAAGATATGTCTTGTGTTGTAGCCAAATACATGGTGTATAGCCCAAAA	3841
QY	3841	atqagaagattattcttcagtagtttgaacacttgtcatcgaatttcagctgcagagctgctc	3900

QY 918 ggcacgcgataacatgcgagggccgagaccacagtgggcctgggctgggtgctgtcttccagca 977
Db 721 GGCACGCCATACATCGAGGCGCGACACAGTGGGCTGGGGCTGGTCTTCACAGCA 780
QY 978 actaacaggcgagcccaacgctgctgtgctatgctccacacatcttcagctcgttgggttt 1037
Db 781 ACTAACAGGGCAGCCCAACGCTGCTGTGCTATGCCCTCCACCATCTTCAGCTCGGTTGGTTT 840
QY 1038 ccatgggggagctcctcagcgcgtgctgctcctctgtgggctgtggcgagtgaaaggtggcagc 1097
Db 841 CCATGGGGATCCTCAGCCGCTGCTGGCCCTCTGTGGGGCTTGGCGCAGTGAAGGTGGCAGC 900
QY 1098 taccctgaccgcgcaatgggctgggtgacogtgcagggccgagggctctgtgtcagctgg 1157
Db 901 TACCCTGACCGCATGGGGCTGCTGGACCTGCTGAGCGCGCAGGGCTCTGTGCTAGCTGG 960
QY 1158 ctgtgcccctcagtgcccctctcgttcagtgagcaatgggctcgtcagcttttgccgtgccat 1217
Db 961 CTGTGCCCTCATGTGCCCTCTCCGCTCAGTGGCATAGGGCTCTGCTAGCTTTGGCGTGCCCAT 1020
QY 1218 ggactcaggcccaagctgtcgtgctgtgcccataatgccacgggcagacagggcctccctgg 1277
Db 1021 GGACTCAGGCCCAAGCTGTGTGGCTGTGCCCAATGSCACCGGGCAGACAGGGCTCCCTGG 1080
QY 1278 agactctggcctgtcgcagagctcctctctacctcccatcccaaggacaaatgagagacca 1337
Db 1081 AGACTCTGGCCTGTGCAGGACTCTCTTACCTTCCCATTTCCAAAGGACCAATGAGGACCA 1140
QY 1338 aaggagagcaatcttgtccactctgaagaaacaaagcccatcccaagctctggagacc 1397
Db 1141 AAGGAGAGCAATCTTGTCCACTGCTTAAGAAACCAAGCCCATCCACAGATCTGGAGACCC 1200
QY 1398 ctgagcccctctcggctggcctgagctcgtgcctcctcctgggcccctctgcgcgctcg 1457
Db 1201 CTCAGCCCTCTCTGGCTGGCCCTGAGCTGTGCCCTCCTCGGGGCCCTCTGCGCGCTGG 1260
QY 1458 ggggcatgcaactgctgcgtggacgcaactgctgtgctgctgctgtcttctgctgctgt 1517
Db 1261 GGGGATGCACTGCTGCTGGACCGACACTGCTGTGCTGATGGTCTTGTGCTAGTGGCTT 1320
QY 1518 ctctcttggtttggggcagtgactgctgtgctcctcagcagatctacccctgtgagat 1577
Db 1321 CTCTTTTGGGTTTGGGCCAGTGACCTGTGGCTTGTCTCTCAGCAGATCTACCCCTGTGGAGAT 1380
QY 1578 acgaggaagagcctctgcgctcttgcacagctctcaactgggcggcccaacctcttcacag 1637
Db 1381 ACGAGGAAGAGCCTTCCGCTTCTGCAACAGCTTCACTGGGCGGCGCAACCTCTTCTATCAG 1440
QY 1638 cctctcctctcctcgtatctcattggccaccatcggtctgtccctggacctctcctgctctacgg 1697
Db 1441 CCTCTCTTCTTCGATCTCATTTGGCACCATCGGCTTGTCTTGGACCTTCTCTCTACGG 1500
QY 1698 actgacgcgtctcctgcgctgggcttcatctataatttctgtcctgaaacaaaaggcca 1757
Db 1501 ACTGACCGCTGTCTCGGCTGGGCTTCACTCTATTATTTCTTCTGAAACAAAAGGCCA 1560
QY 1758 gtcgttggcagagatagaccagagcttccagaagagcgttccacctgagctttgggcca 1817
Db 1561 GTCTGTGGCAGAGATAGACACAGCAGTCTCCAGAGAGACGGTTCACCCCTGAGCTTTGGCCA 1620
QY 1818 caggcagaactcactggcatccctgacagccgcatcgagatctctgcggcctcctcgag 1877
Db 1621 CAGGCAGAACTCCATGGGATCCCGTACACCGCGATCGAGATCTCTGGGGCTCTCTGAGG 1680
QY 1878 aatcgtctgcctggg-aaattctggaactgtggctttggcagaccatctccagcatctctgc 1936
Db 1681 AATCGCTCTCGCTGGAAATCTGGAACGTGTGGCTTGGCAGACCATCTCCACAGATCTCTGC 1740
QY 1937 ttctcagggccacagacaaagtctcagctgggtcttcttgggagtgggccctgcgcccaaa 1996
Db 1741 TTCTTAGGCCCCAGACACAAGTTCAGCTGGTCTTTTGGGAGTGGGCCCTGCCCCCAAA 1800
QY 1997 ggtggttctgcttttgcgtggggtaaaaagagtgaaagtctgagaaatgcccattcttctt 2056

Db 1801 GGTGCTCTGCTTTTGTCTGGGTAAAAAGGATGAAAGTCTGAGAAATGCCCAACTCTTTCATT 1860
QY 2057 ttgggtttcagggccctgaagggtctcttgaggatctagtttcatgctcgtcgttctcccccattg 2116
Db 1861 TTGACTCTCAGGCGCTGGAAGGTTCTTCCAGGATCTAGCTTCTATGCTCAGCTTTCCTCCCATTG 1920
QY 2117 acttgagacatcttttgcagtttttataaagaagaaatattctatgaagtcttttctgcccact 2176
Db 1921 ACTTGACACATCTCTGCAGTATTTATAAGAAGAAATATCTATGAAGTCTTTGTGACCAT 1980
QY 2177 ggatttttttcaaaagaaatctcaggggtaccaaatccgggacaggaaggttttcccccagataca 2236
Db 1981 GGACTTTTCTCAAAAGAAATCTCAAGGGTACCAATCTCGCAGGAAGTCTCTCCCGATATCA 2040
QY 2237 cccctaaatccaaatgagagatacatcttttctaactcttcttcttcttcttcaactgctggaca 2296
Db 2041 CCCCTAAATCCAAATGAGGATATCATCTTTTCTAATCTCTTTTCAACTGGCTGGGACA 2100
QY 2297 ttttcggaagggggaagctctcttttcttcttcttcttcttcttcttcttcttcttcttcttctt 2356
Db 2101 TTTTCGGAAGGGGAAAGTCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2160
QY 2357 ctctctctgttcccaggtgctgctgctcttcttcttcttcttcttcttcttcttcttcttcttctt 2416
Db 2161 CTCATTCTGTGGCCAGGCTGGCTGATCTTGGCTCACTGCAACCTCCACCTCTCTGAGTT 2220
QY 2417 caagcgaattctcctcagcctcagcctccttaagtctggtgattacagggcgctgcccaccaca 2476
Db 2221 CAAGCGATTCTTGTGCTCAGCTCCTTAAGCAGCTGGGACTACAGCGCATGCAACCAT 2280
QY 2477 cccagctaatcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2536
Db 2281 CCCAGCTAATTTATTTTATAGCAGAGATGGGGTTTCTACTGTCTTGGCCAGGCTGGTGTGA 2340
QY 2537 actcctgagctcaagtgatccaccacactcagcctccagctccagagctgctagattacaggcct 2596
Db 2341 ACTCTGAGCTCAAGTGATCCACCACCTCAGCCTCCAGAGTGTAGGATTTACAGGCT 2400
QY 2597 ttgactcttctctctgagtttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2656
Db 2401 TTTGACTCTTTTATCTGAGTCTTTATTTGACCCCTCTAATCTCTTTACCCAGAAATATTATC 2460
QY 2657 ctctcaccagcaactctgactcttcttgcgggagggcctcagttctcagttcttcttcttcttctg 2716
Db 2461 CTTCCACAGCAACTCTGACTCTTTTGGGGAGGCTCAGTTCTACTCTCTTCTCTCTCTCTCTG 2520
QY 2717 tgcctctgctgtaggaatgaccacgggctcagtttcttcttcttcttcttcttcttcttcttctt 2776
Db 2521 TGTCAATTTGCTAGGAATGACCACGGGCTCAGTTTCTCCCATTTTCTAATTTGGGAAGCCT 2580
QY 2777 gtaccaggtcattcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2836
Db 2581 CTACACAGGTCATTTCTTAAAGATTTCTCTGACTCCAGTGAGTGGAAATTTCTAATTTGCTGT 2640
QY 2837 ctgagagctcttccaggaagggtcaggaagggtcttgcggaaaggagatggggttttggagg 2896
Db 2641 CTAGAGAGCTGTCTCCAGGATGGTGACAGATGGCTTTGGCGAAAGAGAGATGGGTTTGGAGG 2700
QY 2897 caacaaactcgtctgtcacaatttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 2956
Db 2701 CCAACAAACCTGCTTGTCTCAATATTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2760
QY 2957 taacaactccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 3016
Db 2761 TAACAACCTCCTGAACCTCAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2820
QY 3017 gtaatttagaccctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 3076
Db 2821 GTATATTAGACCTCTTCTTCTTCTAGGAGGTCCCCAGCTGTGTCTCAGGGCTGGGAAAT 2880
QY 3077 ttctacttactcctcattaccaggtctcctcttcttcttcttcttcttcttcttcttcttcttctt 3136

D	b	2881	TTCTACTTATCTCA	TATACCCAGGTCCTCTCTTTGGACCCCTGTAAAGGCTCAGGGTGAAT	294	
Q	y	3137	caga	tgggggagctgagcaagtagctatgactgcagatca	tgtaaaggaaggggagctgacaag	3196
D	b	2941	CAGATGGGGACTG	AGCAAGTAGCTATGACTGCAGATCA	TGTAAAGAAAGGCACTGACAAG	3000
Q	y	3197	aagctcccagatg	ctgggggagaatgaagagctaaaaatagatccttaggtgctggatgcgttt	3256	
D	b	3001	AAGCTCC	CAGATGCTGGGAGAA	TGAAGAGCTAAATAGATCTTCTAGGTGCTGGATGCTTT	3060
Q	y	3257	gtcatccatg	cggtgcacatcatgggtgctggcagagcccccgaagagactctgcctctcaag	3316	
D	b	3061	GTCTCC	ATGCTGCACATATGGTGCTGGCAGAGCCCCCAAGGACTCTGCCTCTCGAG	3120	
Q	y	3317	tctccctattct	ccattctagatgcttcccttgatc	ccagtgatgctgaggagctggc	3376
D	b	3121	TTCTCC	TATCTCTCCATTTAGTACTTCCCTTGATCCAGTGTATGCTGGAGCTGGC	3180	
Q	y	3377	tttgccaagcttg	tgaagagctggttgctacatttccaggatltttacaagcttggtaaaca	3436	
D	b	3181	TTTGCC	AAAGCTTGAGAGCTGGTGTCTACATTTTTCAGGATTTTTTACAAAGTTGTTAAACA	3240	
Q	y	3437	cagccattata	aaaaattaaa	lgatttaattataatlaagtaaatatacaataaaacaa	3496
D	b	3241	CAGCCAT	TATAAAAATTAAATGATTAAATTTATATTTTANGTAAATTACATTAACAA	3300	
Q	y	3497	aaaaattact	caaaattcattacttaatttactacctgttactattatctgtgcttt	3556	
D	b	3301	AAAAAT	TATATCTCAAAATTCATTA	CTTAATTTTACTACCTGTTACTATTATTCTGTGCTTT	3360
Q	y	3557	tgaaggctattct	atcatagtaactcttttgagacctagggagacacccgcatactctt	3616	
D	b	3361	TGAGGCT	ATTCTACATAGTAAC	TCTTATGGAGACCTAGGGGAGACACCGCATCTCTT	3420
Q	y	3617	cctgattcc	ccactcaatgacatcatgtagtcttcttggtgcttactgagagacacccgagctg	3676	
D	b	3421	CCTGAT	TCCCACTCAATGACATCATGTTAGTCTTTGTTGCTTAAC	TGGCTGTGGGAG	3480
Q	y	3677	tgttttgtat	cacaagattagagaggactacacatcaggcttgatttatgttctgtt	3736	
D	b	3481	TGTTTT	TGTATCACAAAGATTAGAGAGGACTACACATCAGGGCTTGATTATTGTTGTT	3540	
Q	y	3737	gatttctag	actcngaacatgctgataaaaatgcagtaaatgcaaaatlaaaacttaaa	3796	
D	b	3541	GATTTT	CTAGACTTCAGAACATGCTGGATAAAATGTCAGTAATGCAAAATTAACATTTAAA	3600	
Q	y	3797	gtatgctctgt	ttgttagccaatacatatgglgatagcaccacaaaaatggaggattatct	3856	
D	b	3601	GTATG	CTTGTGTGTAGCCCAATACATGGTGTATAGCACCAAAAAATGGAGGATTATCT	3660	
Q	y	3857	tccagtagt	tgaacatgctatccgtttcagctgacagctgctcaaatcatttaagaag	3916	
D	b	3661	TCCAGT	AGTTGAACACTGTCTCCGTTTCAGCTGACACTGCTCAAAATCATTTAAGAAG	3720	
Q	y	3917	agttcgac	attcattcttcttacttttcttctccactagtgtaaacacaaat	3976	
D	b	3721	AGTTCT	GACATTCATTTCTATTTTACTTTTGTCTTCTTCAC	TAGTGTAAACAAAAAT	3780
Q	y	3977	ttcaacc	agcatctatgcogaacctatcccatcttcttcagtgacctagctgtacagttatc	4036	
D	b	3781	TTCA	ACCAGCATTCATGCGGAACCTATACCATTTCTTCAGTGGCTGAGCTACAGTTATC	3840	
Q	y	4037	aggaatttt	tatctgtagtctaatttgtccaaatcatgcccacaaatcgcagtgatagtga	4096	
D	b	3841	AGGGAT	TTTTTATTTGTGTAGTCTAATTTTGTCAAAATCATGGCCAAATCGCAGTGATAGTTGA	3900	
Q	y	4097	ctttgga	tacaaggtttggcaaaaaaaaataatttaacaaaatattcttctaagaatcaat	4156	
D	b	3901	CTTT	GGATACAAAGGTTTGGCAAAAAAAAATATTACAAAAATATCTGTAAAGAAATCAAT	3960	
Q	y	4157	tggctata	tatggaaatttaggataaaagaatattacaataaagaacatttacaataaagaat	4216	
D	b	3961	TGCT	ATATGCAATTTAGATAAAGAAATATTTACAATAAAGAATATTTACAATAAAGAAT	4020	

Qy	4217	ttattattattgtaagttgtgtgcaacaaacataccctttatctctgtgtaaaatttatac	4276
Db	4021	TTATTATTATTGTTAAGTTGTGCAACAACACATACCCCTTTATCTCTGTGTAATTTATAC	4080
Qy	4277	acacaaaaattacaaaagattctgtgaagaatttaattggctcatatggaatttaggataga	4336
Db	4081	ACACAAAAATTACAAAAGATTCTGTGAAGATTAAATGGCTATATGGAATTTAGGATAGA	4140
Qy	4337	atatattacaataaagagattattacaat	4363
Db	4141	ATATTTACAAATAAAGAGTATTTTACAAT	4167

RESULT 3

HS28H201
 LOCUS
 DEFINITION
 PRI 21-NOV-2000
 HS28H201 4126 bp mRNA
 Novel human gene mapping to chromosome 20, similar to membrane transporters.
 ALI37188
 ALI37188.3 GI:11322734
 .
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4126)
 Stavrides,G.S., Hashim,Y., Huckle,E.J. and Deloukas,P.
 Direct Submission
 Submitted (24-JAN-2000) E-mail contact: humquerry@sanger.ac.uk
 On Nov 23, 2000 this sequence version replaced gi:11065679.
 This cDNA sequence was assembled from public domain ESTs and single pass sequencing reads from expressed DNA templates, aligned to the genomic DNA sequence from the bacterial clone 28H20 (AL031055). The EST sequences listed match this sequence with an identity of at least 95% between the coordinates shown.
 Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr20/ Sanger Centre name :
 dJ28H20.C20.1.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

exon	1	12
CDS	number=1	9..1634
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="CAB69822.2"	
	/db_xref="GI:11065680"	
	/translation="MGHSPVPLPCASVSLGLGTFGYELAVISGALLPLQIDFGLSC LEGEFLVGLLGLALLASLVGGFLDICYGRKQAILGSLNVLLAGSLTGLAGSLAWLV PGWRIMFGFAISLSSMACCIYVSELVGPQRGVLSVLYEAGITVGLISYALNYALAGT LGLFRARDNNRGRTTVGLGLVLFQOLTGPNVCLYASTIFS VGFHGSGSVAVLASVNP GAVKYAATLTAMGLVDNRGRALLUAGCALMALSVSGILGVFAVPMDSGSPCLAVPN ATGQTGLGDSGLQDSSLPPIPTNEDQREPI LSTAKTKT KPHRPSGDSPAPRIALFS SALPGPLPARGHALLRWLTALLCMVFSAFSGFGPVTWLVLSEIYVPEITRGRAFAF CNSFNAAANFI FLSLFDLIGTIGLSWTFLLVGLTAVLGLGFIYLVFVPEIKQGS LAEI DOOQKRRFT"LSFGCHRONSTGIPYSRIEISAAS"	
exon	13	1296
	/number=2	
exon	1297	..1419
	/number=3	
	1338	..1643
misc_feature	/note="matches EST AA313045"	
exon	1420	..1555
	/number=4	
exon	1556	..1426
	/number=5	
misc_feature	2095	..2137

		/note="matches EST AI753293 from clone HBMSC_cr08b03 complement(join(3694..3898,3867..4123))
misc_feature		/note="matches EST AI097288 from clone IMAGE:1707216"
misc_feature		complement(3727..4126)
misc_feature		/note="matches EST AI292321 from clone IMAGE:1894742"
misc_feature		complement(join(3737..3898,3867..4126))
misc_feature		/note="matches EST AI753932 from clone HBMSC_cr16c08"
misc_feature		complement(join(3774..3898,3867..4126))
misc_feature		/note="matches EST AI277131 from clone IMAGE:1893578"
misc_feature		complement(3793..4126)
misc_feature		/note="matches EST N93207 from clone IMAGE:304871"
misc_feature		complement(3795..4124)
misc_feature		/note="matches EST AI289525 from clone IMAGE:1932441"
misc_feature		complement(join(3795..3887,3877..4017,4005..4126))
misc_feature		/note="matches EST AW103571 from clone IMAGE:2614034"
misc_feature		complement(3867..4125)
misc_feature		/note="matches EST AW166863 from clone IMAGE:2634465"
misc_feature		complement(3867..4123)
misc_feature		/note="matches EST AA532534 from clone IMAGE:986291"
misc_feature		complement(3867..4126)
misc_feature		/note="matches EST AI753418 from clone HBMSC_cr10b04"
misc_feature		complement(3945..4126)
BASE COUNT	914 a 1066 c 979 g 1167 t	
ORIGIN		
Query Match 92.5%; Score 4064.4; DB 9; Length 4126;		
Best Local Similarity 99.2%; Pred. No. 0;		
Matches 4094; Conservative 0; Mismatches 31; Indels 1; Gaps		
Qy	242	cgtcgccatgggcactcccacctgctccctgctttgtgtgcctctgtgttccttgcctgcagc 361
Db	1	CGCTCGGCATGGCCACTCCCCACCTCCTCGTCCCTTTGTGTGCCCTGTGTGCTTTTCCTGG 60
Qy	302	gtggcctgaccttggtltaagaactggcagtcatactaggtygccctgtgcacactgcagc 361
Db	61	GTGGCTGACCTTTTGTTATGAAGTGGCAGTCATATCAGGTGCCCTGCTGCCACTGCAGC 120
Qy	362	ttagctttggctaaagtgcgttggagcaggaattcctggtgggcagcctgtcctcctgggg 421
Db	121	TTAGCTTTGGGCTAAGCTGCTTGAGCAGGAGTTCTCTGCTGGCAGCCTGCTCCTGGGG 180
Qy	422	ctctcctgcctcctcgttgttgggtgggtctctcatatgactgctataggcaggagaagcca 481
Db	181	CTCTCCTCGCCTCCCTGGTTTGGTGGCTTCTCATTTGACTTGCTATGGCAGGAACAAGCCA 240
Qy	482	tctcgggagcaacttgggtgctgctgagcaggcagcctgaccctgggcctggtcgtgtccc 541
Db	241	TCCTCGGAGAACAACTTGGTGCTGCTGTCAGCAGCACCTGACCCTTGGGCTGGGTGTCCC 3000
Qy	542	tggcctggctggttccttgggcgcgcgtggtgtgcttcgcctttccctcctcctcctgg 601
Db	301	TGGCCTGGCTGGTTCCTGGGCGGGCTGTGGTGGCTTCGCGCATTTCCCTCTCCTCCATGG 3606
Qy	602	cttgcgtatctacgtgtgcagactggttggggccacgcagcgggagtgctcgtgttccc 661
Db	361	CTTGCCTGTATCTACGTGTACAGCTGTGGGGCCACGCAGCGGGGAGTGTGTGTCTCC 4200
Qy	662	tctatgaggcaggcatacacgtgggcatcctgctctcctatgcccccaactatgcactgg 721
Db	421	TCTATGAGGCAGGCATCACCGTGGGCATCTCTGCTCTCTATGCCCCCTCAACTATGCAC 4800
Qy	722	ctgttacccccctggggaatggaggcacatgttcggctggggccactgcacctgctcctgc 781
Db	481	CTGTGTACCCCTCGGGGATGAGAGGCACATGTCGGCTGGGCGCCACTGCACCTGCTCCTGC 5400
Qy	782	aatccctcagcctcctctcctcctcctcctgctggtatcacagatgagactgacaacacaa 841
Db	541	AATCCCTCAGCCTCCTCTTCTCCTCCCTGCTGGTACAGATGAGACTGCAACACACAGGACC 6000
Qy	842	tcattccactccaggagggtgaggcccccaagctgggcccggggagggcacaggtactcct 901

QY 3061 ccaggccctgggaaatttctactctatccctcattaccagggtccctccttttgagccctgta 3120
DB 2821 CAGGGCCCTGGGAAATTTCTACTTATCTCTATCTATCCAGGTCCCTTTGGACCCCTGTA 2880
QY 3121 aagggcaggggtcaatcagatggggagctgagcaagtagctagctagctagctagctagctaa 3180
DB 2881 AAGGGTCAGGGTCAATCAGATGGGGGACGTGAGCAAGTAGCTAGCTAGCTAGCTAGCTAA 2940
QY 3181 ggaaggagctgacaagaagctcccgatgctggggagagtagagagctaaaaatagatcct 3240
DB 2941 GGAAGGAGCTGACAAAGAAGCTCCAGAGTCTGGGAGAGTAAGAGCTAAAATAGATCCT 3000
QY 3241 aggtgctgagctgttttctccatccatgctgctgacacatggtgctgagagagcccccagg 3300
DB 3001 AGGTGCTGGATGCTTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 3060
QY 3301 actctggcctcctcaggtcagtg 3360
DB 3061 ACTCTGGCCCTCTCGAGTCTCTCTATCTCTATCTCTATCTCTATCTCTATCTCTATCT 3120
QY 3361 atgtgctgagctggcttctccagctgttgagagctggttctctacattttcaggaatttt 3420
DB 3121 ATGTGCTGGAGCTGGCTTTGCAAGCTTTGTGAGAGCTGGTTGCTATCTATCTATCTATCT 3180
QY 3421 tacaagctgggtgaacacagcattatatacaaaattaaatgaatttaatttaatttaagta 3480
DB 3181 TACAAGTGGTGAACACAGCCATTTATAAAATTAATTAATTAATTAATTAATTAATTAAT 3240
QY 3481 aattacattaaaaacaaaaattatactcaaaatttacttaatttactactcctgta 3540
DB 3241 AATTACATTAAACAAAAAATTATCTCAAAATTCATTACTTAATTTACTACTCTGTTA 3300
QY 3541 ctattctctgtctgttttgaggtctattctacatagtaactctctatggagacctaggag 3600
DB 3301 CTATTATCTGTCTTTGAGGCTATTTCTACATAGTAACTCTATGGAGACCTTAGGGGAG 3360
QY 3601 acaccgcatctctctctgattcccccactcaatgaatcatgtagtttgggtgctt 3660
DB 3361 ACACCGCATCTCTCTGATTCCCCACTCAATGACATCATGTTTGGTGTGCTT 3420
QY 3661 aactgctgaggagtgctttgtatcacaaagattagagaggactacacacaggct 3720
DB 3421 AACTGCTGTGGGAGTGTTTGTGTATCAAAAGATTAGAGAGGACTACATCAGGGCT 3480
QY 3721 tgatttattgtttgtgatttcttagacttcagaaactgctggaataaaatgctagtaag 3780
DB 3481 TGATTTATTCTTGTCTGATTCTTAGACTTCAGAACATGCTGATATAAATGTCAGTAATG 3540
QY 3781 caaatbaaacttbaaagatgctgtttgttagccaaatacataggtgtatagcccaaaaa 3840
DB 3541 CAAATTAACCTTAAAGTATGCTTGTGTAGCCCAATACATGGTGTATAGCACCAAAAA 3600
QY 3841 atggaggattattctccagtagtgaacactgctacccgttccagctgacagctgctc 3900
DB 3601 ATGGAGGATTATTCTTCCAGTAGTGAACACTGTCATCCGTTTTCAGCTGACAGCTGCTC 3660
QY 3901 aaatcttaagaaggagctgacattcttctctctctctctctctctctctctctc 3960
DB 3661 AAATCATTTAAGAAGGAGTCTGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3720
QY 3961 tagtgaacacaaatctcaaccagcattctgcgaacactatccattctcctcagtgcc 4020
DB 3721 TAGTGTAAACAAAAATTTCAACAGCATTCATGCCGAACCTATACCCATCTCTCAGTGCC 3780
QY 4021 tagctgacagttatcaggagatttttattctgtagtctaaatttctgtaaaatcatggccaaa 4080
DB 3781 TAGCTGTACAGTTATCAGGATTTTATTGTGTAGTCTAATTTTGTCAATCATGAGCCAAA 3840
QY 4081 tcgcagtgatgatttgacttgatgacaggtttggcaaaaaaaattatttaacaaata 4140
DB 3841 TCGCAGTGATGATTGACTTTGGATACAGGTTTGGCAAAAAAATAATTAACAAAAATA 3900
QY 4141 ttctgtaagaatcaatttggctatagtggaatttaggataaagaataatttacaataaagaat 4200

DB 3901 TTTCTGTAAGAATCAATTTGTCTATATGGAATTTAGATATAAGAATATTTTCAATAAGAAT 3960
QY 4201 atttacaataaagagtttattattatttatttatttatttatttatttatttatttatttatt 4260
DB 3961 ATTTCATAAAGAGTTTATTATTATTATTATTATTATTATTATTATTATTATTATTATTATT 4020
QY 4261 tctgtaaaatttatac 4320
DB 4021 TCTGTAATAATTTATACACACAAAAAATTAACAAAGATTTCTGTAAGAATTAATTTGGCTATA 4080
QY 4321 tggaaattaggatagaattatttacaataaagaagtagtatttacaataaaa 4366
DB 4081 TCGAATTTAGATAGATAATTTTACAATAAAGAGTATTATTACAATAAA 4126

RESULT 4
HS28H20 127418 bp DNA PRI 19-APR-2001
LOCUS Human DNA sequence from clone RP1-28H20 on chromosome 20q13.1
DEFINITION Contains the SIC2A10 gene encoding a solute carrier family 2 (facilitated glucose transporter) member 10, the 5' end of a novel gene, ESTs, STSs, GSSs and three CpG islands, complete sequence.
ACCESSION AL031055.1 GI:4375937
VERSION 1
KEYWORDS HTG; CpG island; SIC2A10.
SOURCE human.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 127418)
TITLE Ramsay, H.
JOURNAL Direct Submission
COMMENT Submitted (18-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Mar 7, 1999 this sequence version replaced gi:4056528.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20
This sequence is the entire insert of clone RP1-28H20 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP1-28H20 is from the library RP1-1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.
Location/Qualifiers
1. 127418
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="q13.1"
/clone="RP1-28H20"
/clone_1lb="RPC1-1"
complement(join(<1. 562,2463. 3110))

FEATURES
source

mRNA

```

/gene="dJ101A2.2"
/note="match: cDNAs: Em:AK013049 Em:AK011574
match: ESTs: Em:AA233726 Em:T87841 Em:AA756018 Em:AA113277
Em:AA985248 Em:AA682750 Em:AA461487 Em:AI028588
Em:AA659710 Em:AI149891 Em:AA579641 Em:AA516077
Em:AA460132 Em:AA483490 Em:AA046575"
/evidence=not_experimental
/product="dJ28H20.2 (novel protein)"
complement(1..3110)
/gene="dJ101A2.2"
18..494
/note="match: STS: Em:HS28H20T"
complement(join(84..562,2463..2745))
/gene="dJ101A2.2"
/note="supported by FGENSEH and GENSCAN
continues in Em:AL133520 as dJ101A2.2"
/codon_start=1
/evidence=not_experimental
/product="dJ28H20.2 (novel protein)"
/db_xref="GI:9588402"
/db_xref="SPREMBL:O9NO66"
/translation="MAAARATTPADGPEAPAEALAAARSSRFSLGSLVKGCAE
ARVFRGFCGRAAVIKHRFPKGYRHPALRLGRRTVQEARALLRCRAGISAPVVF
FVDYASNCILYMEIEGTVTRDVIQSTMETEKTPQGLSNLAKTIQVLIARHDEDLIH
GDLTTSMLLKPPLEQLNIVLIDFGLSFTSALPEDKGVLDLYVLEKAFLSHPNTEVF
EALFSYSSSSKKARPLVKLLDEVLRGRKRMVG"
867..1145
/note="AluX repeat: matches 1..298 of consensus"
1251..1353
/note="MIR repeat: matches 50..151 of consensus"
1672..1797
/note="MIR repeat: matches 8..135 of consensus"
1855..2079
/note="MIR repeat: matches 33..262 of consensus"
2155..2248
/note="MIR repeat: matches 33..135 of consensus"
2384..3150
/note="CpG island"
/evidence=not_experimental
2812..2871
/note="3 copies 20 mer 85% conserved"
complement(2876..3110)
/gene="dJ101A2.2"
3865..3962
/note="match: STS: Em:G15621"
/note="MER21B repeat: matches 299..391 of consensus"
3963..5655
/note="L1MA8 repeat: matches 4532..6270 of consensus"
5669..5895
/note="L1ME repeat: matches 5501..5739 of consensus"
6332..6548
/note="L2 repeat: matches 2524..2746 of consensus"
6626..6915
/note="AluX repeat: matches 3..290 of consensus"
6918..6980
/note="MER9A repeat: matches 1..62 of consensus"
7023..7322
/note="AluJo repeat: matches 1..312 of consensus"
7323..7417
/note="L2 repeat: matches 2389..2513 of consensus"
7457..7623
/note="MIR repeat: matches 93..262 of consensus"
7864..7934
/note="MLT1I repeat: matches 311..382 of consensus"
8046..8240
/note="MLT1J repeat: matches 9..211 of consensus"
8347..8647
/note="AluSq repeat: matches 1..300 of consensus"
8685..8793
/note="MLT1J repeat: matches 271..366 of consensus"
9095..9221
/note="MIR repeat: matches 102..242 of consensus"

```

```

repeat_region
9647..9776
/note="L2 repeat: matches 2170..2303 of consensus"
9793..10028
/note="AluJo repeat: matches 74..306 of consensus"
10038..10206
/note="LTR16C repeat: matches 126..316 of consensus"
10631..10690
/note="MLT1D repeat: matches 445..505 of consensus"
10691..10995
/note="AluJb repeat: matches 1..295 of consensus"
10996..11200
/note="MLT1D repeat: matches 248..445 of consensus"
11232..11634
/note="MSTB repeat: matches 2..409 of consensus"
11670..11761
/note="MLT1D repeat: matches 128..192 of consensus"
11762..12062
/note="AluSx repeat: matches 1..293 of consensus"
12063..12180
/note="MLT1D repeat: matches 1..128 of consensus"
complement(12181..12344)
/note="match: GSS: Em:AQ190762"
13011..13129
/note="MLT1J repeat: matches 252..368 of consensus"
13210..13392
/note="MER20 repeat: matches 1..215 of consensus"
13397..13442
/note="MIR repeat: matches 104..146 of consensus"
13660..13783
/note="LTR33 repeat: matches 330..452 of consensus"
13894..14254
/note="L2 repeat: matches 2255..2619 of consensus"
14305..14435
/note="L2 repeat: matches 2617..2748 of consensus"
14446..14525
/note="L2 repeat: matches 20 mer 76% conserved"
15476..15752
/note="AluJb repeat: matches 2..278 of consensus"
15774..15952
/note="MER58A repeat: matches 42..220 of consensus"
16190..16304
/note="L2 repeat: matches 2657..2744 of consensus"
16305..16599
/note="AluJb repeat: matches 1..296 of consensus"
16600..16635
/note="L2 repeat: matches 2617..2657 of consensus"
16641..16803
/note="L1MA10 repeat: matches 6162..6284 of consensus"
16804..17190
/note="MER57B repeat: matches 1..403 of consensus"
17198..17297
/note="5 copies 20 mer 71% conserved"
17204..17297
/note="47 copies 2 mer ta 70% conserved"
17559..17868
/note="AluSx repeat: matches 1..312 of consensus"
17939..18240
/note="LTR26 repeat: matches 246..553 of consensus"
complement(18333..18408)
/note="match: STS: Em:G42885"
18451..18615
/note="FAM repeat: matches 126..41 of consensus"
19258..19336
/note="MER37-internal repeat: matches 7168..7244 of
consensus"
19343..19485
/note="MER53 repeat: matches 2..180 of consensus"
19721..20190
/note="LTR8 repeat: matches 251..691 of consensus"
20238..20493
/note="LTR8 repeat: matches 1..256 of consensus"
20704..20978

```

Query Match 57.6%; Score 2531.4; DB 9; Length 127418; Best Local Similarity 98.8%; Pred. No. 0; Matches 2561; Conservative 0; Mismatches 31; Indels 1; Gaps 1;			
QY 1796	ggttcacccctgagcttttggccacacagggcagaaactcacttggaatcccggtacacagccgcatcg	1855	
DB 47086	GGTTCAACCCCTGAGCTTTGGCCACAGGAGCAACTCCACTGGCATCCGGTACAGCCGCATCG	47145	
QY 1856	agatctctcgccctctctgaggaatccgtctgcctgg-aattctggaactgtggtcttgg	1914	
DB 47146	AGATCTCTGGGGCTCTCTGAGGAATCCGCTGCGCTGGAATTTCTGGAACCTGCGCTTGG	47205	
QY 1915	cagaccatctccagcatcctgtctctctcctagggcccccagagacacagttccagctgctcttt	1974	
DB 47206	CAGACCATCTCCAGCATCTCTCTCTCTAGGCCCCAGAGCACAAGTTCCAGCTGGCTCTTT	47265	
QY 1975	ggagtgggcccttgcccccaaaagtggttctgcttctgctggggttaaaaggaaggaagtt	2034	
DB 47266	GGGAGTGGCCCTTGCCCCAAAGGTGGTCTGCTTTGCTGGGGTAAAGGATGAAGTC	47325	
QY 2035	tgagaatgcccattctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt	2094	
DB 47326	TGAGAATGCCCAACTCTTCAATTTTCACTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG	47385	
QY 2095	tcatgctcggtttcccccatttgacttgacacatttttgcagtttttataagaagaatttc	2154	
DB 47386	TCATGCTCAGTTTCCCATTTGACTTGCACATCTCTGCAATTTTATAAGAAGAAATTC	47445	
QY 2155	tatgaagctcttcttgcctcattgatcttcttcaagaatctcaggggtaccacatccggg	2214	
DB 47446	TATGAAGCTCTTGTGGCACCATGAGACTTTCTCAAGAAATCTCAAGGATACCAATCTCG	47505	
QY 2215	caggaggttttcccgatatacccccataatcccaaatgaggatatacatcttcttctaatc	2274	
DB 47506	CAGGAAGCTCTCCCGATATCACCCCTAAATCCAAATGAGGATATCATCTTTTCTAATCT	47565	
QY 2275	cttttttcaactggctgggaacatttttgggaagggggaagttcttcttcttcttcttct	2334	
DB 47566	CTTTTCTCAACTGGCTGGGACATTTTCGGAAGGGGGAAGCTCTCTTTTACTCTTATCA	47625	
QY 2335	tcttttttttttggagtggaatctcattctctgttgcacaggttggcctgatcttggctcac	2394	
DB 47626	TTTTTTTTTTTTTGAAGTGGAGTCTCATTTCTGTGCCCCAGGCTGGCCTGATCTTGGCTCAC	47685	
QY 2395	tgcaactccactctcctgggttcaagcgtctctcgtcctcagcctccctaaagtagctggg	2454	
DB 47686	TGCAACCTCCACCTCTCTGAGTTCAAGCCATTTCTGTGCTCTCAGCTTCTTAAGCAGCTGG	47745	
QY 2455	attacagggcgtgcccacacacacacagctaattttttttagcagagatggggtttcact	2514	
DB 47746	ACTACAGGCCGATGCAACCATACCCAGCTAATTTATTTTACGAGAGATGGGGTTTCACT	47805	
QY 2515	ggttggccaggtctggtgaaactcctgagctcagtgatccacacacacacacacacacccc	2574	
DB 47806	GTCTTGCCAGGCTGGTCTGCAACTCCCTGAGCTCAAGTCAAGTCAATCCACCCAGCTCCC	47865	
QY 2575	agagtgctagattacagggccttttgactcttcttcttcttcttcttcttcttcttcttct	2634	
DB 47866	AGAGTGTAGATTTACAGGCTTTTGTACTCTTTTATCTGAGTTTATTTACCCCTCTAAT	47925	
QY 2635	tctcttaccagaaatatttcttcttccacagcaactctgactcttctgaggggagggcctca	2694	
DB 47926	TCCTTTACCCAGAAATATTTATCTTTCACAGCAACTCTGACTCTTTTGGGGGAGGCTCA	47985	
QY 2695	gttctagctcttggctgctggtgctcattctgtaggaatgaccacggcctcagttctcc	2754	
DB 47986	GTTCTAGCTCTTGGTCTGCTGCTCATTTGCTGAGGAATGACCAAGGCTCAGTTTCC	48045	
QY 2755	ccatttgtataatgggaagcctgtaccaggtctcttcttcttcttcttcttcttcttcttct	2814	
DB 48046	CCATTTGTATAATGGGAAGCTGTACCAGGTCAATTTTCAAGATTTTCTCTGACTCCAGTG	48105	
QY 2815	agctggaattcttaaatgctgctgcttaggagctgtctccagggatggtgcaggatggtcttgc	2874	
DB 48106	AGCTGAATTTCTAATCTCTGCTTAGGAGCTGTCTCCAGGATGGTGCAGATGGCTTTC	48165	
QY 2875	ggaagagagatgggttggagggccacacacacacacacacacacacacacacacacacac	2934	
DB 48166	GGAAGAGAGATGGGTGGAGGCCCAACAACTGCTGTGCAATATTTGCCCTTTCCTCTTG	48225	
QY 2935	gagcccttgaacttgagtaataacacacacacacacacacacacacacacacacacacac	2994	
DB 48226	GCAGCCCTTCAACTTGAGTAATAAACAACCTCCCTCAACCTCAGTTTCTCTATCTGCAGAA	48285	
QY 2995	tggggataatagtcccccaggggtatatattagacccctgttcccttccagggaggggtcccca	3054	
DB 48286	TGGGGATAATATATGCTCCAGGGTATATTTAGACCTGTCTTCTTTCAGAGGAGTCCCA	48345	
QY 3055	gctgtccagggcctgggaaatttcttcttcttcttcttcttcttcttcttcttcttcttct	3114	
DB 48346	GTGGTCCAGGGCTGGGAAATTTCTACTTATCTCTATTTACCAGCTCCCTCTTGGAC	48405	
QY 3115	cctgtaaaggttcagggatgaatcagatggggagcagcagatgctggggagaaagaaagata	3174	
DB 48406	CCTGTAAAGGGTCAAGGTGAATCAGATGGGGACTGAGCAAGTAGTATGACTGCGACATC	48465	
QY 3175	atgtaaaggaaggaactgacaaagagcctccagatgctggggagaaagaaagataaata	3234	
DB 48466	ATGTAAAGGAAGGACTGACAAAGAGCTCCAGATGCTGGGGAGATGAGAGGCTTAAATA	48525	
QY 3235	gactcaggtgctgagatgcttcttctcactcagctgagcagatgctgggctgagagagccc	3294	
DB 48526	GATCCTAGGTGCTGGATGCTTTGTCTCATCCATCGCTGCACATATGGGTGCTGGCAGAGCCC	48585	
QY 3295	caaaggaactcgtgctcctcgtgctcctcctcctcctcctcctcctcctcctcctcctcct	3354	
DB 48586	CCAAGGACTCTGGCTCTCGAGTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	48645	
QY 3355	caagtgatgctgctgagagctgcttgcacagcttgcagagctggtgctgacatcttctcag	3414	
DB 48646	CCAGTGTGCTGGAGCTGGCTTTGCCAAGCTGTGAGAGCTGGTTGTACATTTTTCAG	48705	
QY 3415	gatttttacaagttggttaaac	3474	
DB 48706	GATTTTACAGTTGGTAAACACAGCCATTTATAAATAATTAATGATTTAATTTATAAT	48765	
QY 3475	taagtaaatcacatataaac	3534	
DB 48766	TAAGTAAATTTACATTAACAAACAAACAAATTTACTCAAAATTTCAATTAATTTACTAC	48825	
QY 3535	ctgttactattatctgtcttcttcttcttcttcttcttcttcttcttcttcttcttcttct	3594	
DB 48826	CTGTTACTATTATCTGCTTTTGGAGCTATTTCTACATAGTAACCTTTATGGAGACCTA	48885	
QY 3595	ggggagacacccgcacatctcttcttcttcttcttcttcttcttcttcttcttcttcttct	3654	
DB 48886	GGGACACACCCGCATCTCTCTCTGATTTCCCACTCAATGACATGATGTTAGTCTTTGG	48945	
QY 3655	tgtcttaactggtctggtgggagtggttcttcttcttcttcttcttcttcttcttcttctt	3714	
DB 48946	TTGCTTAACTGGCTGGGGAGTGTTTTGTATCAAAAGATTTAGAGAGACATACATC	49005	
QY 3715	agggcttgatttattgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct	3774	
DB 49006	AGGGCTTGATTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	49065	
QY 3775	gtaatgcacaaatttaaaactttaaagtgcttcttcttcttcttcttcttcttcttcttctt	3834	
DB 49066	GTAATGCAAAATTAACCTTAAAGATGCTTGTGTTGTAGCCCAATACATGCTGATATACAC	49125	
QY 3835	caaaaaatggagggatattcttcttcttcttcttcttcttcttcttcttcttcttcttctt	3894	
DB 49126	CAAAAAATGGAGGATTTATTTCTCCAGTAGTTGAACACTGTCTATCGTTTTCAGCTGACAG	49185	
QY 3895	ctgtcacaatcatttaagaagaggttcttgcacattcttcttcttcttcttcttcttcttct	3954	

Db 49186 CTGCTCAATCATTAAGACGAGTCTGACATTTTCATTTTACATTTTGTCTT 49245
 QY 3955 cctcactagtgtaacaaattttcaacagcattcatgccaactatccattcttc 4014
 Db 49246 CTTCACTAGTGTAAACAAAAATTTCAACGAGCATTCATGCGAACCTATACCCATCTTC 49305
 QY 4015 agtgccctagctgacagttatcagggatttttattcgttagtctatatttgcataatcag 4074
 Db 49306 AGTGCCTAGCTGTACAGTATATCAGGATTTTATTGTAGTCTAATTTGTCAAAATCATG 49365
 QY 4075 gcaaatcgcagtgatagtgtagcttgatgatacaagggttggcaaaaaaaataattatac 4134
 Db 49366 GCCAAATCGAGTGATAGTGTACATTTGATGATCAAGGTTTGCAAAAAAATAATTAAAC 49425
 QY 4135 aaaaattcttgtaagaatcaattggtctatgcatggaatttagatgaagaataatttacaata 4194
 Db 49426 AAAATATCTGTAGAAATCAATTTCTATATGGAATTTAGGATAAAGAAATATTACAATA 49485
 QY 4195 aagaatatttacaataaagaggtttatttatttatttatttatttatttatttatttattt 4254
 Db 49486 AAGAATATTACAAATAAAGAGTTATTATTATTATTATTATTATTATTATTATTATTATT 49545
 QY 4255 ttattctctgtaaaatttatac 4314
 Db 49546 TTTATCTCTGTAAATTTTATACACACACACACACACACACACACACACACACACACAC 49605
 QY 4315 gctataggaatttagatagataatttacaataaagaggtatttacaataaagaggttctg 4374
 Db 49606 GCTATATGAATTTAGATAGAAATTTTACAATAAAGAGTATTTACAATAAAGAGTTGT 49665
 QY 4375 tattattgttaa 4387
 Db 49666 TATTATTGTAAA 49678

RESULT 5

AL591064 200125 bp DNA HTG 03-MAY-2001
 LOCUS Mus musculus chromosome 2 clone RP23-395E18, *** SEQUENCING IN
 DEFINITION PROGRESS ***, in unordered pieces.
 ACCESSION AL591064
 VERSION AL591064.3 GI:13990651
 KEYWORDS HTG; HTGS-PHASE1.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 200125)
 Sims.S.
 Direct Submission
 Submitted (02-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On May 7, 2001 this sequence version replaced gi:13990284.

Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Center project name: Project Information
 Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: Plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 195932 bases at least Q40
 Consensus quality: 196830 bases at least Q30
 Consensus quality: 197490 bases at least Q20
 Insert size: 198525; sum-of-contigs
 Insert size: 206466; 5.9% error; agarose-fp
 Quality coverage: 8.60x in Q20 bases; sum-of-contigs Quality
 coverage: 8.32x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES
 source
 1..200125
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-395E18"
 /clone_lib="RPCI-23"
 1..4470
 /note="assembly fragment:01119
 fragment_chain:1
 clone_end:T7
 vector_side:left"
 4571..16275
 /note="assembly fragment:03729
 fragment_chain:1"
 16376..33874
 /note="assembly fragment:02569
 fragment_chain:1"
 33975..46195
 /note="assembly fragment:05190
 fragment_chain:1"
 46296..57940
 /note="assembly fragment:09914
 fragment_chain:1"
 58041..62650
 /note="assembly fragment:03121
 fragment_chain:1"
 62751..71940
 /note="assembly fragment:01509
 fragment_chain:1"
 72041..80555
 /note="assembly fragment:02606
 fragment_chain:1"
 80656..85958
 /note="assembly fragment:01353
 fragment_chain:1"
 86059..113228
 /note="assembly fragment:04226
 fragment_chain:1"
 113229..122676
 /note="assembly fragment:00791
 fragment_chain:1"
 122777..153882
 /note="assembly fragment:03653
 fragment_chain:1"
 153983..162005
 /note="assembly fragment:02332
 fragment_chain:1"
 162106..184926
 /note="assembly fragment:04610
 fragment_chain:1"
 185027..190517
 /note="assembly fragment:01620
 fragment_chain:1"
 190618..196019
 /note="assembly fragment:00987
 fragment_chain:2"
 196120..200125
 /note="assembly fragment:00065
 fragment_chain:2
 clone_end:SP6
 vector_side:right"
 BASE COUNT 50402 a 48944 c 49953 g 49224 t 1502 others
 ORIGIN

Query Match 18.1%; Score 795; DB 2; Length 200125;
 Best Local Similarity 76.6%; Pred. No. 9,6e-140;
 Matches 991; Conservative 0; Mismatches 290; Indels 12; Gaps 1;

Qy	253	ggcaactcccacactgtctcgtccttggtagcctctgtgtcttctgtgggtggcctgacc	312
Db	163943	gGCCTTGCCCAAGCTGCTCCTCCTGCTCTGTGCTCTGTGCTGCTTTGCTGGGGGGCTGACC	164002
Qy	313	ttggcttatgaactggagtcacatacaggtgcctctgctgcacatgcagcttgactttggg	372
Db	164003	TTGGCATATGAGCTGGCCGTCATATCTGGTGGCCCTCCTACCACTCCAGCTGAACCTTCGGG	164062
Qy	373	ctaagctgcttggagcagaggttcctctgggtgggcagcctgctctctgggggtctctcctgcgc	432
Db	164063	CTGAGTTGCTCCGAACAGAGAGCTCCGGTGGCGAGTCTGCTTTGGGTGCTCTCCTCGCC	164122
Qy	433	tcctctgggttggtagcttctcattgactgctatggcaggaagaagcaccatcctctcgggaagc	492
Db	164123	TCCCTTGTGGGGGGTTCCTCATTTGACTGCTATGGCAGGACAGAGCATCCTTAGGGAGC	164182
Qy	493	aacttggtagctgtggcagcagccttgaccctggcctggcctggctgggttctccctggcctggctg	552
Db	164183	AATGCGGTGCTGCTGGCTGGCAGCCTGATTCTGGGTCTGGCCAGCTCCCTGCCCCGTGGCTG	164242
Qy	553	gtccttggccgcgcgtgtgttggcttgcgcatttccctctccctccatggcttgcctgtgtatc	612
Db	164243	CTCCCTGGGCCGTTTATCTGTGTGGCTTTGGCATCTCTCTGTCTTCATGGCTTGTGTATTC	164302
Qy	613	tacgtgtcagagctggttggggccagcgagcgaggagtgctggtgtccctctatgagaca	672
Db	164303	TACGTGTCCAGAAATTGCTGGGGCGCGCAGAGGGGTGCTGTGGTGTCCCTCTACAGAGTG	164362
Qy	673	ggcatcacctgtgggatacctgctctcctatgcccctcaactatgcactggcttggtaaccoc	732
Db	164363	GGCATCACTGTGGGCATTCTGTCTTACGGGCTCAACTACGTTCCTGGCTGGCAGCCCT	164422
Qy	733	tgggtagtgagggacacatgttcggctgggcgcactgacactgctgtctctgcgaatcctcagc	792
Db	164423	TGGGCTGGAGGCACATGTTGGCTGGCAGCTGCACCTGCTCTCCTACAGTCACTCAGC	164482
Qy	793	ctcctcttctcctctgtctgttaacagtagactgcaacacacagaagacctcactcccactc	852
Db	164483	CTCTTTTCTCCTCTGCTGTGCAGAGGGCACAGAGCCCCCAAGACCTTATCCCCACTC	164542
Qy	853	caggaggttgaggcccccaagcttgggcccggggaggccacggtactcctttcttggaactc	912
Db	164543	CAGGGAAGGGGACAAAGCAAAACCGGGCTTGGTGAAGGCACAGTACAGCTTTTCTGGACCTC	164602
Qy	913	ttcaggggcacgcgataaacatgcgaggccggacacacagtgggcctggggctggtgctcttc	972
Db	164603	TTTCAAGGCCACGACAGCGCATGTGAGCCCGACCTGTATGTGGGGCTTGGGGCTGTGCTGT	164662
Qy	973	cagcaactaaacaggcagccccacgttgccttgcctccactcctcactctcagctccgtt	1032
Db	164663	CAGCAGCTCACAGGACAGGCCCAATGTGCTGTATTACGGCTCTACCATCTTTTCGCTCCGTT	164722
Qy	1033	ggtttccatgggggatacctcagcgcgttgcctctcttggggcttggcgagtcgaagtgc	1092
Db	164723	GGCTTCCACGGGGGCTCCTCAGCTGTGCTGGCTTCTGTGTGGGGCTTGGCAGGTTGAAGGTG	164782
Qy	1093	gcagctacactgacgcgcatagggcttggtagccgtgcagggccgcagggctctgttgcata	1152
Db	164783	GCTGCCACCTTGTTCGCCACCGGGGCTGGTGACCGTGCAGGGCGCAGAGTCTCTCTGCTC	164842
Qy	1153	gctggctgtgcctcctatggcccttgtccgttcagtgggcatagggcctcgttcgctgt	1212
Db	164843	TTTGGATGGCTCTCATGGCTTATCGGTGAGCGGCATAGGCCCTGGTTCAGCTTTGCTGTG	164902
Qy	1213	cccatggactcagggcccaagctgtctctggctgtgcccgaatgccaccgggcagacagccctc	1272
Db	164903	TCCCTGGACTCTGGACCCAGCTGCCTGGCCACATCCAAATGCCAGCCAGCAGGTGATCTG	164962
Qy	1273	cctggagactcttggcctgtgtgaggactcctctctactcctcccatcccaaggacccaatgag	1332
Db	164963	CTTGAAGCTCGGGTTTGTCTGTGGCTTCTATCTCTGCGGCAGTGCCTACACACCAATGGG	165022

[illegible]

FEATURES

```

source      location/Qualifiers
1. 243075
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-90N15"
/clone_lib="RPCI-23"
1. 45622
/note="assembly_fragment:03140"
clone_end:17
vector_side:left"
45723. 53487
/note="assembly_fragment:00309"
fragment_chain:1"
misc_feature
misc_feature

```

RESULT	7
LOCUS	G31518/c
DEFINITION	G31518 411 bp DNA
ACCESSION	human STS SHGC-44522, sequence tagged site.
VERSION	G31518
KEYWORDS	G31518.1 GI:1912168 STS; STS sequence; primer; sequence tagged site.

LOCUS G31518 411 bp DNA STS
 DEFINITION human STS SHCC-44522, sequence tagged site.
 ACCESSION G31518
 VERSION G31518.1 GI:1912168
 KEYWORDS STS sequence; primer; sequence tagged site.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 411)
 AUTHORS Myers, R.M.
 JOURNAL Unpublished (1997)
 COMMENT
 Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu
 Primer A: ACTATCACTGGGATTTGGCC
 Primer B: GCCGACCTATACCCATTCT
 STS size: 100
 PCR Profile:
 Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600
 Protocol:
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul
 Buffer:
 MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.3
 Prepared with primer pairs provided by Sandoz, derived from N93207
 -- Washington University/Merck EST sequence.
 FEATURES
 source
 location/Qualifiers
 i. .411
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="20"
 STS
 primer_bind 308..407
 primer_bind complement(388..407)
 BASE COUNT 134 a 60 c 49 g 164 t 4 others
 ORIGIN
 Query Match 8.8%; Score 384.6; DB 11; Length 411;
 Best Local Similarity 97.8%; Pred. No. 1e-62;
 Matches 398; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 Qy 3989 tcatgccgaacctatbaccattcttcagtcgctagctgtacagttatcagggaattttat 4048
 Db 411 TCATGCCGACCTATACCCATTCTTCAGGCGCTAGCTGTACA-NTATCAGGCGATTTTAT 353
 Qy 4049 tcgtagctcaattttgtcaaatcatcgcccaaatcgagtgatagttgactttggatacaa 4108
 Db 352 TCTTAGCTCAATTTTGTCAATCATGCGCAATCGCAGTGTAGTTGACTTTGGATACAA 293
 Qy 4109 ggtttggcaaaaaaaattatacaaaaattattctgtaagaatcaattggctatatgga 4168
 Db 292 GGTTCGCAAAAAAATAATATTAAACNAATATTCTGTGAAGCAATCAATTTGGCTATATGGA 233
 Qy 4169 atttaggataaagaattattcacataaagaattatttacaataaagaagttattattttt 4228
 Db 232 ATTTAGGATAAAGATATTTTNCATATAAGATATATTTCATATAAGAGTTTATTATT 173

Oy 4229 gaaagtgtgtgcaaaaacatacccttttattctctgttaaaattttacacacaaaaatta 4288
 Db 172 GTAAGTTGTGAGCAACAACATACCCCTTTATCTCTGTGTAATTTTATACACACAAAATTA 113
 Qy 4289 acaaaagattctgtagaataattggctatatatgaatttaggatagaataattacaata 4348
 Db 112 ACAAAAGATTCTCTAAGAAATAATTGGCTATATGGAATTTAGGATAGATAATTTTACAATA 53
 Qy 4349 aagagtatttacaataaagaagatttggctattatttggtaaaaaaataa 4395
 Db 52 AAGAGTATTTACAATAAAGAGTTTGTATTATTATTGTTGTAATAAANAANA 6
 RESULT 8
 AX150151/c 385 bp DNA PAT 08-JUN-2001
 LOCUS Sequence 126 from Patent WO0136685.
 DEFINITION AX150151
 ACCESSION AX150151
 VERSION AX150151.1 GI:14348179
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 385)
 AUTHORS Kroes, R.A., Moskal, J.R. and Yamamoto, H.
 TITLE Differential gene expression in cancer
 JOURNAL Patent: WO 0136685-A 126 25-MAY-2001;
 NYXIS NeuroTherapies, Inc. (US)
 FEATURES
 source
 location/Qualifiers
 i. .385
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 127 a 61 c 43 g 154 t
 ORIGIN
 Query Match 7.2%; Score 317.4; DB 6; Length 385;
 Best Local Similarity 97.5%; Pred. No. 4.4e-50;
 Matches 344; Conservative 0; Mismatches 6; Indels 3; Gaps 2;
 Qy 4017 tgcctagctgtacagttatcagggaattttttatctgtagtcttaattttgtcaaatatggc 4076
 Db 373 TGCCGAGCTGTACAGTTATCAGGGATTTTATTTTGTAGTCTAATTTTGTCAATCATGGC 314
 Qy 4077 caaatcgagtgatagttgactttggatacaaggtttggc-aaaaaaaaataataaca 4135
 Db 313 CAATCGCGGTGATAGTTGATTTTGGATACAGGTTTGGCAAAAAAATAATTTAACA 254
 Qy 4136 aaatattctgaagaatcaattggctataggaatttaggataaagaatatttacaataa 4195
 Db 253 AATATCTGTGAAGATCAATTGTCTATATGGAATTTAGGATAAAGA--ATTACATAA 196
 Qy 4196 agaatatttacaataaagaagttttattatttttgaagttgtgtgcaacaataccct 4255
 Db 195 AGAATATTTACATAAAGAGTTTATTATTATTATTGTAAGTTGTGTCACAAACATACCT 136
 Qy 4256 ttatctctgaattttatcac 4315
 Db 135 TTATCTCTCTAAAAATTATACACACAAAAATTAACAAAAAGATTCTGTAGAATTAATTGG 76
 Qy 4316 ctatatggaatttagatagaataatttacaataaagaagttatttacaataaaga 4368
 Db 75 CTATATGGAATTTAGGATAGAATATTACATAAAGAGTATTTACAATAAAGAGTATTTACAATAAANA 23
 RESULT 9
 AB056798 4521 bp mRNA PRI 14-MAR-2001
 LOCUS Macaca fascicularis brain cDNA clone:Qf1A-11110, full insert
 DEFINITION sequence.
 ACCESSION AB056798

[illegible]

[illegible]

repeat_region	11741..11947 /note="Charlie1b repeat: matches 1..215 of consensus"
repeat_region	12018..12207 /note="Charlie1b repeat: matches 322..521 of consensus"
repeat_region	14943..15073 /note="MR58 repeat: matches 139..284 of consensus"
repeat_region	15362..15635 /note="MR7A repeat: matches 39..346 of consensus"
repeat_region	15774..16087 /note="AlusX repeat: matches 1..302 of consensus"
repeat_region	16088..16222 /note="AlusG/x repeat: matches 1..137 of consensus"
misc_feature	16386..17708 /note="CpG Island" /evidence=not_experimental
repeat_region	17282..17387 /note="MIR repeat: matches 112..208 of consensus"
repeat_region	17762..17882 /note="L1MC4 repeat: matches 7854..7976 of consensus"
repeat_region	17943..18263 /note="AluJb repeat: matches 3..311 of consensus"
repeat_region	18271..18684 /note="L1MC4 repeat: matches 7345..7809 of consensus"
repeat_region	20635..20776 /note="L2 repeat: matches 2304..2443 of consensus"
repeat_region	20927..21034 /note="3 copies 36 mer 75% conserved"
repeat_region	21543..21695 /note="MR68B repeat: matches 1..156 of consensus"
repeat_region	21764..22038 /note="MR68A repeat: matches 266..549 of consensus"
repeat_region	22221..22329 /note="Charlie1b repeat: matches 263..392 of consensus"
repeat_region	22664..22952 /note="AlusX repeat: matches 1..302 of consensus"
repeat_region	22965..23809 /note="L1MB5 repeat: matches 5352..6164 of consensus"
repeat_region	24286..24423 /note="3 copies 46 mer 77% conserved"
misc_feature	24892..25059 /note="match: STS: Em:G23729"
misc_feature	25837..26065 /note="match: STS: Em:G20380"
misc_feature	26057..26057 /note="match: STS: Em:G30039"
repeat_region	28726..29023 /note="match: STS: Em:G05845"
repeat_region	31138..31437 /note="AlusX repeat: matches 1..298 of consensus"
misc_feature	31483..31946 /note="match: STS: Em:G56295 match: GSS: Em:AQ321323"
repeat_region	32645..32790 /note="L1P4 repeat: matches 5271..5405 of consensus"
repeat_region	32912..33200 /note="AlusG repeat: matches 1..291 of consensus"
repeat_region	33290..33459 /note="L1MB5 repeat: matches 5128..5310 of consensus"
repeat_region	33480..33797 /note="AlusX repeat: matches 1..312 of consensus"
repeat_region	34065..34520 /note="MLT1D repeat: matches 23..505 of consensus"
repeat_region	34658..34711 /note="27 copies 2 mer cc 75% conserved"
repeat_region	34664..34713 /note="10 copies 5 mer cctcc 92% conserved"
repeat_region	34715..34854 /note="L1M4 repeat: matches 4443..4600 of consensus"
repeat_region	34863..35158 /note="AluY repeat: matches 1..295 of consensus"
repeat_region	35160..35334

[illegible]

RESULT	15
AL356577/c	
LOCUS	AL356577 85378 bp DNA HTG
DEFINITION	Homo sapiens chromosome 11 clone RP1-6104, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION	AL356577

